Recursive Estimation of User Intent from Noninvasive Electroencephalography using Discriminative Models

Niklas Smedemark-Margulies[1], Basak Celik[2], Tales Imbiriba[2], Aziz Kocanaogullari[2,3], Deniz Erdogmus[2]

[1]Khoury College of Computer Science, Northeastern University
[2]Department of Electrical and Computer Engineering, Northeastern University
[3]Analog Devices

Task

- Goal: restore communication using non-invasive EEG
- Rapid serial visual presentation (RSVP)
  - Query subject with a sequence of possible symbols
  - Measure responses
  - Consider context and responses, update estimated symbol probabilities

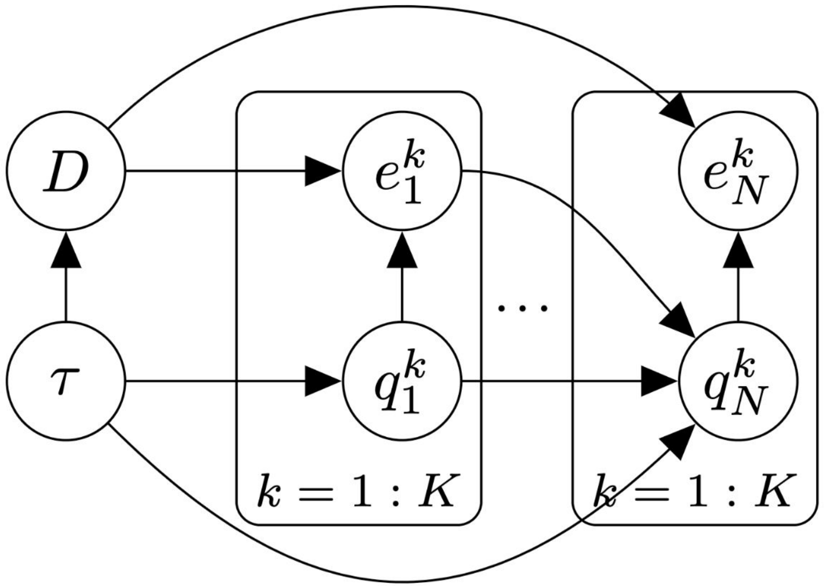

Fig. 1: Probabilistic graphical model for RSVP typing task after N rounds of query and response, with K symbols per query. D - user's target symbol. τ - previously typed text. q_N^k and e_N^k - symbols and EEG responses during N^{th} query.

Recursive Bayesian Update Rule

Presented symbol: α
Other symbols: β
Binary label: ℓ

1. $\gamma_N^k(D{=}\alpha) = \frac{p(\ell{=}+|e)}{p(\ell{=}+)}\pi_N^{k-1}(D{=}\alpha)$

2. $\gamma_N^k(D{=}\beta) = \frac{p(\ell{=}-|e)}{p(\ell{=}-)}\pi_N^{k-1}(D{=}\beta).$

3. $\pi_N^k(D{=}\alpha) = \frac{\gamma_N^k(D{=}\alpha)}{\sum_{d\in(D)} \gamma_N^k(d)}$

4. $\pi_N^k(D{=}\beta) = \frac{\gamma_N^k(D{=}\beta)}{\sum_{d\in(D)} \gamma_N^k(d)}$

Recursive Bayesian Update Rule

Presented symbol: α
Other symbols: β
Binary label: ℓ

1. $\gamma_N^k(D{=}\alpha) = \dfrac{p(\ell{=}+|e)}{p(\ell{=}+)}\pi_N^{k-1}(D{=}\alpha)$

2. $\gamma_N^k(D{=}\beta) = \dfrac{p(\ell{=}-|e)}{p(\ell{=}-)}\pi_N^{k-1}(D{=}\beta).$

3. $\pi_N^k(D{=}\alpha) = \dfrac{\gamma_N^k(D{=}\alpha)}{\sum_{d\in(D)} \gamma_N^k(d)}$

4. $\pi_N^k(D{=}\beta) = \dfrac{\gamma_N^k(D{=}\beta)}{\sum_{d\in(D)} \gamma_N^k(d)}$

COGNITIVE SYSTEMS LAB.

Recursive Bayesian Update Rule

Presented symbol: α
Other symbols: β
Binary label: ℓ

1. $\gamma_N^k(D{=}\alpha) = \dfrac{p(\ell{=}+|e)}{p(\ell{=}+)}\pi_N^{k-1}(D{=}\alpha)$

2. $\gamma_N^k(D{=}\beta) = \dfrac{p(\ell{=}-|e)}{p(\ell{=}-)}\pi_N^{k-1}(D{=}\beta).$

3. $\pi_N^k(D{=}\alpha) = \dfrac{\gamma_N^k(D{=}\alpha)}{\sum_{d\in(D)} \gamma_N^k(d)}$

4. $\pi_N^k(D{=}\beta) = \dfrac{\gamma_N^k(D{=}\beta)}{\sum_{d\in(D)} \gamma_N^k(d)}$

Recursive Bayesian Update Rule

Presented symbol: α
Other symbols: β
Binary label: ℓ

1. $\gamma_N^k(D{=}\alpha) = \frac{p(\ell{=}+|e)}{p(\ell{=}+)} \pi_N^{k-1}(D{=}\alpha)$
2. $\gamma_N^k(D{=}\beta) = \frac{p(\ell{=}-|e)}{p(\ell{=}-)} \pi_N^{k-1}(D{=}\beta).$
3. $\pi_N^k(D{=}\alpha) = \frac{\gamma_N^k(D{=}\alpha)}{\sum_{d \in (D)} \gamma_N^k(d)}$
4. $\pi_N^k(D{=}\beta) = \frac{\gamma_N^k(D{=}\beta)}{\sum_{d \in (D)} \gamma_N^k(d)}$

COGNITIVE SYSTEMS LAB.

Experimental Evaluation

RSVP Benchmark Dataset (Zhang et al, 2020[*])

- 64 subjects, >1M binary trials total

Models evaluated:

- Discriminative neural nets
- Discriminative classic models (LDA, LogR)
- Baseline generative models

Metrics:

- Balanced accuracy
- Information Transfer Rate (ITR) in simulated typing (see Alg 1)

Algorithm 1: Estimating ITR via simulated typing. Note that the likelihood L predicted by the model at each step can be either $p(e|\ell)$ or $p(\ell|e)/p(\ell)$, as described in Sec. 2.3.

Input: Trained model $f(\cdot)$, Pos. and Neg. Test Data $\mathcal{X}^+, \mathcal{X}^-$, Iterations T, Symbols per query K, Attempts per symbol N, Alphabet size A, Decision threshold δ,

Output: ITR

$C \leftarrow 0$ `// correct count`
for $t \leftarrow 1 : T$ **do** `// target symbols`
    $\pi_0 \leftarrow (\frac{1}{A}, \ldots, \frac{1}{A})$ `// unif symbol prior`
    for $n \leftarrow 1 : N$ **do** `// chances to update`
        `// sample query symbols`
        $\{q_i\}_{i=1}^K \sim \pi_{n-1}$
        `// sample matching data`
        $\{x_i \sim \mathcal{X}^+$ **if** $q_i = t$ **else** $x_i \sim \mathcal{X}^-\}_{i=1}^K$
        $L \leftarrow f(\{x_i, q_i\})$ `// model likelihoods`
        Calc. π_n from π_{n-1} and L `// Eq.9-11`
        `// see if target was typed`
        ind, val $\leftarrow \arg\max(\pi_n), \max(\pi_n)$
        if ind=t and val $\geq \delta$ **then** $C \leftarrow C + 1$ and **break**
return ITR$(A, C/T)$

*https://doi.org/10.3389/fnins.2020.568000

COGNITIVE SYSTEMS LAB.

Results

Table 1: Balanced Accuracy and Information Transfer Rate (ITR) for Discriminative (Disc) and Generative (Gen) Models. The discriminative strategy yield models with higher balanced accuracy and information transfer rates. Entries show mean and standard deviation across 5 random train/test splits. Control models use the discriminative strategy but always assign high probability to a fixed class. See Sec. 3.6 for ITR calculation.

Strategy	Model	Balanced Acc	ITR
Disc	LogR	0.730 ± 0.001	0.817 ± 0.047
Disc	EEGNet	0.745 ± 0.003	0.930 ± 0.050
Disc	1D CNN	**0.782 ± 0.005**	**1.103 ± 0.047**
Disc	2D CNN	**0.779 ± 0.004**	**1.153 ± 0.068**
Gen	LDA (Emp Prior)	0.509 ± 0.000	0.678 ± 0.077
Gen	LDA (Unif Prior)	0.687 ± 0.003	
Gen	LogR (Emp Prior)	0.500 ± 0.000	0.218 ± 0.022
Gen	LogR (Unif Prior)	0.694 ± 0.002	
Control	Always Class 0	0.500 ± 0.000	0.000 ± 0.000
Control	Always Class 1	0.500 ± 0.000	0.000 ± 0.000

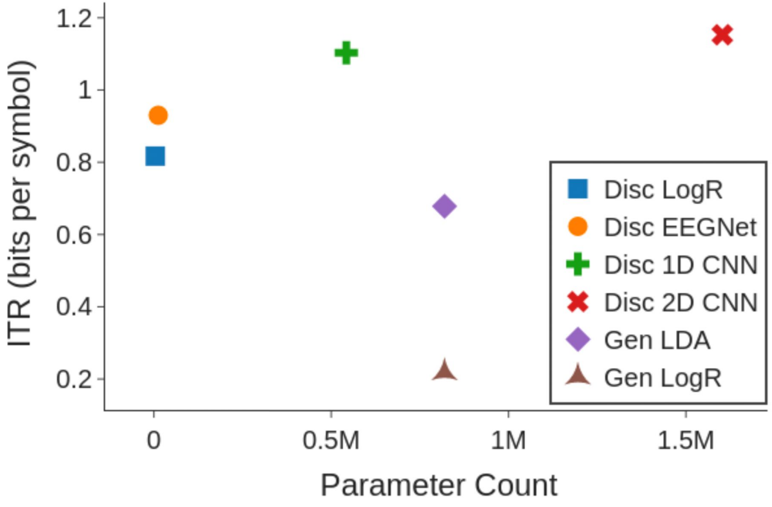

Fig. 2: Information Transfer Rate vs Model Size. Discriminative (Disc) models outperform generative (Gen) models across a wide range of sizes. Among Disc models, performance increases with model size.

Thanks

Code for experiments: https://github.com/nik-sm/bci-disc-models

Code for dataset: https://github.com/nik-sm/thu-rsvp-dataset

Full paper: https://arxiv.org/pdf/2211.02630.pdf